

# **Groupsize's influence on punishment in public good game.**

*Freeriding in social groups; are there any difference in smaller versus larger groups?*

**Brigt Erland Nersveen**

**17.08.2007**

---

**Department of Economics  
University of Oslo**

## Preface

This is my master thesis after three years at the Department of Economics, University of Oslo.

First of all I will thank my mentor, Karine Nyborg, for being patient with me and guiding me in the right direction whenever a problem evolved.

My second thoughts and appreciations go to Karen Hauge and Sigurd Brandanger for valuable teamwork, both during planning and execution of the treatments. Our teamwork is in fact one of the factors making all our experiments running smoothly.

Additional: Frisch center for taking me in to their environment, giving me a workstation and a place to write and study.

Professor Karl Erik Moene for making my thesis possible.

The experiment was programmed and conducted with the software z-Tree (Fischbacher 2007)

# Content

Preface .....	ii
Content .....	iii
Summary .....	1
1 Introduction .....	3
1.1 Generality .....	3
1.2 Rough overview of the experiment .....	4
1.3 Public good .....	6
1.4 Freeride .....	6
1.5 Short on Z-Tree .....	7
1.6 Motivation .....	7
2 Background .....	8
2.1 The beginning of public good games .....	8
2.2 Introduction of punishment .....	8
2.3 Punishment in public good games .....	9
3.1 Game technicalities .....	11
3.1.1 Set up .....	11
3.2 Rules .....	12
3.2.1 Introduction .....	12
3.2.2 Basic rules and information .....	13
3.2.2 Setup .....	13
3.2.3 Questionnaire .....	19
3.3 More on game setup .....	19
3.3.1 Contribution phase .....	19
3.3.2 Punishment phase .....	20
3.4 Treatment differences .....	22
4 Theory .....	24
4.1 Homo economicus .....	24
4.2 Other theories .....	25
4.2.1 Inequity aversion .....	25
4.2.2 Reciprocity .....	27
4.2.2 Altruism .....	28
5 Results .....	30
5.1 Problems .....	30
5.1.1 Router .....	31
5.1.2 Program .....	31
5.1 Data summary .....	32
5.1.1 Contribution and punishment .....	32
5.1.2 Individual punishment received .....	33
5.2 Why this “simple” regression? .....	35
5.2.1 Critique .....	35
6 Summary and conclusion .....	36
7 Appendix .....	37
7.1 Instructions for the participants in Norwegian .....	37
7.1.2 Treatment 16-1 .....	37
7.1.3 Treatment 16-4 .....	40
8 References .....	43

## Summary

This thesis is concerned with a topic in behavioral economics that is a relatively new and fast growing area of study. An area close tied up with psychology since we are looking into how people behave under certain scenarios or settings.

This thesis is built up by motivational and theoretical discussion, before looking at the data collected during two experiments using Z-Tree software by Urs Fischbacher. For regression I used Stata 9.0.

The experimental setting I used were a variant of a so-called public good game with punishment. In this game the participant contribute to a common good and get something in return from this good. It is constructed such as to give a conflict between individual and group interests. After the contribution it is possible to punish other players on the background of their and everybody else's contribution behavior.

The motivational section contains short history of public good game and behavioral economics as well. One important aspect of my experiment is the punishment. Since I'm doing a public good game where participants contribute to a common good, I have opened up for punishment by revealing contributions of individuals. This revealing is of course anonymous, but participants can target punishment based on these contributions. This will be rigorously explained in main body of text.

My main motivation for this thesis is to look into if groupsize has an effect on how much people punish each other. Therefore, I constructed two treatments; one where participants could punish all other participants (16-1) and one where they could only punish three other assigned to same group (16-4). The contribution group was equal in both treatments, i.e. all contributed to the same common good/pool. This way I ensured that there was only one difference between the two treatments.

My hypothesis is as follows:

“For any given individual contribution by individual  $i$ , this individual  $i$  is punished less by others in the 16-1 treatment than in the 16-4 treatment.”

Then there is a chapter of how we conducted the experiment. We held the experiment in Norwegian, but all the rules and instructions are reproduced in English. In the appendix, I have enclosed a Norwegian version of the instructions as presented to the participants. All participants got a handout of the instructions in addition to an oral presentation by reading the exact content by my assistant. At the end of this chapter, a more mathematical presentation of the game is given and a clarifying part of the difference between the two treatments as mentioned above.

A theoretical chapter explaining alternative behavioral traits comes after this. Here I look at Homo economicus, as a selfish and rational actor with preferences for only maximum economic outcome. I compare this with other preferences that can explain why people contribute and punish in this game setting.

The result chapter is somewhat limited, since one of my treatments failed. However, I did a simple linear regression to see if there was any effect of groupsize on punishment. The lack of complete dataset resulted in inconclusive outcome, but I got an indication that my hypothesis may be right. This can be a motivation for further studies into this problem; disintegrating of individual punishment in increasing groupsize. I also included some discussion of what went wrong as to enlighten those who may intend to use the same setup as I did.

# 1 Introduction

You are in a group who need to cooperate to accomplish a certain goal, and one or more of you are not making adequate effort. How do you react? If you are the lazy one, how do the others react? How do you feel in any case? Do you think the size of the group matters for how you feel or react?

The purpose of this thesis is to study individual behavior in public good games with sanctioning options. More specifically I want to explore whether the level of punishment from individuals is dependent on group size all else given, i.e. whether individual members of smaller groups punish more on an individual level, than they do in larger groups.

## 1.1 Generality

No matter how you live your life, you are bound to interact with other people. There may be some extreme cases where this is not true, but to live a normal life some interaction is not avoidable. You have to shop your groceries and other essentials, most of us have to work, many emphasize spending their spare time with friends, and you probably end up with a family at one point. By all this interaction, we are trying to cooperate with each other. In this thesis, I am focusing on how people react to different exertions toward a common good. Contribution and punishment set in a controllable environment.

Sometimes you interact with few people, sometimes with many. This can very well influence the behavior of individuals: how you react to them and vice versa. Take a study group for example. Here you can choose your degree of preparation. You can prepare fully and contribute a lot to the discussion, or you can prepare nil and just at no cost harvest information from the discussion (not taken account of time spent during discussion). It is of course possible to prepare somewhat between fully and none at all, but the main point is how the others react to your preparation, and how is your reaction towards the others. Let us say that the study group of three has one lazy member. You may sanction him by facial expression (low cost) or directly by telling him what you think (high cost). Imagine the same setting, just in a bigger group. Then it may be more likely that each individual sanctions less believing others will step up and punish instead. This could then lead to a disintegrating of individual responsibility for punishment.

I will in this thesis look into some aspects regarding punishment behavior with difference in groupsize. Before commenting any further on the above examples, I present an overview of the experiment.

## ***1.2 Rough overview of the experiment***

To collect data on behavioral traits resembling situations pictured in the preceding chapter I conducted an economic experiment. I did this economic experiment at a newly acquired computer lab<sup>i</sup> consisting of laptop computers using the software Z-Tree. Z-Tree is a programming environment for setting up numerous types of experiments. For example, dictator/ultimatum games, public good games, auction games and so forth. Urs Fischbacher, a well renowned economic experimentalist, is the man behind this software<sup>ii</sup>.

Even though the study group example presented above does not involve monetary sanctions, this is the main incentive in most economic experiments. In the real world punishments often is a mixture of psychological and physiological expressions, maybe to a greater extent than economic. With bodily expressions, we instantly think of psychology, and the case is that experimental economics is not that different (Rabin, 2002). Economists often use terms as rational, self-interested, stable utility functions, etc. However, in real life, humans do not always adhere to these “economic rules”. Much of behavioral economics has focused on exactly prying into such problems, by experimenting in laboratory or field setting. Contrary to psychologists, economists need measurable incentives, and a natural choice is monetary rewards and punishments.

In chapter 3 I will go through the experiment in detail, but for motivational and background it is in place to give a simple outline. The main goal of the experiment is to see if large group constellations punish less than smaller ones at an individual level. To examine this I am doing a so-called public good game with punishment. This is a setup where participants can choose to contribute a share of their endowment to a common project. The sum of all contributions is multiplied by a factor (larger than one) and then divided equally between the participants. In the next phase participants have the opportunity, on the basis of information of contribution and outcome, to punish each other.

---

<sup>i</sup> Oeconlab at the University of Oslo

<sup>ii</sup> For more information see: <http://www.iew.unizh.ch/ztree/index.php>

I do this twice with different subject pools. The contribution phase is equal in both treatments. The only difference is group size in the punishment phase. In the first treatment each subject gets information of all the participants' contributions, but can only punish three exogenously selected other participants. In the second treatment each get the same information as in treatment 1, but here they can punish every other participant.

This way I can examine if larger groups punish less than smaller ones at an individual level, given all else equal.

While capturing the main intuitive idea of my hypothesis, the studygroup example is not entirely in line with my experiment. In the study group example, groupsize in contribution and punishment was equal. When forming a study group, you will most certainly interact with all the group members, both in contribution (individual preparation for the group) and in punishment (you meet all in the group at a set time for discussion), i.e. the two phases have equal groupsize. A correct model for examining that exact scenario would be to have two treatments where contribution- and punishment-groups are equal inside treatments, but different between treatments. In my design contribution group is equal in both treatments, but the punishment group differs in size.

One other thought enlightening this is that if you have a study group of for example 10. When you meet the seating is organized so you can only interact fully with those close to you, then you have an ad hoc scenario, where group size differs in punishment, but are equal in contribution.

The experimental goal can be of great interest in many ways, from social policies to firm behavior. In firms with an extended team based means of productions, an analytical approach to team size and structure could be of utter importance. Another probably important factor is the flow of information inside teams and between teams. In that connection we can view this thesis as trying to explain how people react to this flow of information about individual effort. My experimental setting builds upon information of other participant's contribution. After the contribution phase (more detailed walkthrough of the experiment in chapter 3) every individual participant get an overview of the others contribution and payoff from that phase. There will not be any analysis of informational flow, but rather the reaction to it.



In social context, there may also be a difference between small clusters of inhabitants compared to large ones. If you for example witnessed a crime or immoral act, are the reaction and action you take dependent of the size your social environment? How should the government then deal with this distinction in their pursuit for equality for all citizens?

### **1.3 Public good**

In my thesis, I use a game type called public good experiment (Ledyard, 1995). In such games, participants get an endowment of and then are asked to decide how much of this they will contribute to a common good. All contributions to the common good are multiplied by a factor and divided equally among the participants.

In economics, the standard definition of a public good is that it is non-rival and non-excludable. Non-rival is when one person's consumption not lowers another person's consumption possibilities, and non-excludable is that there is not possible to shut a person out from consumption (Stiglitz, Walsh, 2002). The standard public good game I am doing here is not fully in accordance to this definition. In chapter 3.1 we see that the public good (all contributions multiplied by 2) is rival, that should indicate that the good in fact is a "common pool resource". Then again, a "common pool resource" experiment is a term used for a slightly different game type. Therefore, we should stick to the well-grounded terms naming this a public good game.

For an example of common pool resource game read "Covenants with and without sword: Self governance is possible" (Ostrom, 1992).

### **1.4 Freeride**

One may think of freeriding, vaguely, as consuming more than a fair share of a resource, or contribute less than a fair share towards production. I believe everyone has experienced a freeriding problem in his or her daily life. It can be housework, voluntary communal work, carrying furniture together with another person, attending a study group etc. In this thesis, we have two types of free riding, one in the contribution- and one in the punishment phase.

Thus, the larger group there is in the punishment phase, the more people can sanction the same person. If punishers focus on this, they may think that their individual effect of punishment is smaller the more people there is to punish. At the same time, there are more people to sanction. We have two effects that pull in opposite directions, one involving freeriding, and one of social responsibility concerning punishing defectors. If my hypothesis is correct, freeriding dominates and disintegrates the responsibility for punishment.

### **1.5 Short on Z-Tree**

Since this was the first time anyone had done any programming in Z-Tree at the University of Oslo, I used quite a good deal of time learning the programming language. A bit more time than planned, but I believe this was important to get a robust script and intuitive screen layout. The time used at programming reduced the chances of doing more radical program changes, if some treatments failed. In such a case I would be certain that there were little chance that this were caused by programming error. Nevertheless, problems still emerged as stated below.

The first time I lost all the data. It seemed that one of the Z-Leafs (participant's computer) went down, making the computer not respond at all. After a while, this made the Z-Tree (server program at the experiment leader's computer) to go into no response. Somehow, all data were lost even though I believe Z-Tree writes and stores data between each period.

The second time similar problems appeared and it stopped after three completed periods. Luckily, these data were stored and saved. We wanted to run the questionnaire, and to do that we need to end a treatment successfully. Therefore, after restarting all Z-Leafs (one on each computer) and restarting the experimenter computer from scratch, two more successful periods were completed. A discussion of this failure is in chapter 5.1 concluding the problem is probably not in the Z-Tree program in itself.

### **1.6 Motivation**

The above chapters give a rough introduction of my motivation and experimental method. Somehow, I find it intriguing how people behave in social settings and experimental economics does this at point. Leaving out most noise that you find in environments that are more open and focus our attention at a specific modeled situation can give us valuable data to look into rather dynamic situations. I try to look into how people react to differences in

contribution dependent on how large social group they are a member of. Giving me my hypothesis:

Hypothesis: Small groups lead to more punishment than bigger ones in a public good game with punishment, all else equal.

## **2 Background**

Public good games and experimental economics as a whole seem to be on the move in the community of economists. There emerges new interesting papers continually, taking up a broad specter of exciting problems.

### ***2.1 The beginning of public good games***

Why do people contribute towards a common good, when they as individuals will be better off by freeriding? Several papers look into this. One early paper is with public good game by Marwell and Ames (Marwell and Ames, 1981). Their angle was the freeriding problem, the tragedy of commons, presented by Hardin (Hardin, 1968). They wanted to test this theoretical paper with experimental evidence. Their major finding is that in public good game participants contribute more than predicted by theory assuming income-maximizing players. Their participants contributed between 40% and 60%, which should indicate that people do not act entirely as assumed by this theory predicting no contribution. One interesting point from this paper is that people who studied economics contributed the least, around 20%. For my thesis, the relevant aspect is that people do contribute to give the participants something to react to in the next phase, the punishment phase.

### ***2.2 Introduction of punishment***

Preference for punishment behavior was first stated with one-shot ultimatum games (Güth, Schmittberger, Schwarz 1982), (Camerer, Thaler 1995). This is a game where we have two participants. One will be asked to divide an amount of money between himself (proposer) and the other person (receiver). The receiver can accept the offer and the payoff will be as proposed by the proposer. If, however, the receiver rejects the offer, the outcome will be zero for both of them. Even though every division can be supported as equilibrium, the subgame perfection will give us that it is likely that the proposer get all (or almost all) of the surplus.

The reason for this is that an income-maximizing receiver will accept any positive amount, no matter how small. Knowing this, the proposer proposes the least positive amount he is capable of. Contradictory to this, we find that small offers usually are rejected, and the reason as stated by participants post experiment is that they punish greedy proposers (Pillutla, 1996).

### ***2.3 Punishment in public good games***

We have confirmed above that people do want to punish others who they feel do an unfair act. I will now show you that this also is the case for public good games with punishment. In a public good game, there is a lot of room to freeride as stated by Marwell and Ames (1981). How do the participants react to such freeriding when they have the opportunity to punish each other? There is a wide selection of papers covering this, and other problems using public good games with punishment. Fehr and Gächter did a basic setup of a public good game with punishment (Fehr and Gächter, 2000). In this paper, they showed that people punished freeriders even though this was a costly act. This contradicts the behavior of income-maximizing individuals, who under such a context should not punish at all.

The Fehr and Gächter experiment consisted of several periods. This can be a critical point, especially if there is not enough participants to run a total stranger set up, i.e. people never meet the same player twice in the same group during the experiment. Let me clarify. If people have a possibility to end up in the same group with another individual more than once, there are stronger incentives to punish as to build reputation. By this, you are telling the other player that you are willing to sacrifice substantial income, just to punish aberrant behavior. By doing this you hope to signal a credible act of punishing, as to make the other person contribute more in the following periods. However, in the final period there should then be no contribution even when such strategic incentives are present, but the case is that there is punishment even in the last round. A method to avoid this problem of reputation is to have enough participants to do a total stranger set up. This means that no one will ever meet the same player twice during the treatment, and thereby no strategic incentives to punish. Even with this setup, people punish, indicating that there are other factors influencing people's utility than just maximizing-income (Fehr, Gächter, 2000).

In my experiment, it is costly to punish. You should think that punishment behavior changes with how cost and effect relate to each other. With effect I mean how much a punishment

point assigned, reduces the income of the receiver. One experiment is looking into how changes in cost and effect of punishment alter punishment behavior of the participants (Nikiforakis, 2005). They conclude that you need a somewhat effective punishment before the contribution goes up in repetitive periods. Let's say that punishment factor (amount of punishment you inflict by spending one token) is  $p$ . They conclude that for  $p = 0, 1, 2$ , the contribution is declining in repetitions of the game. In addition, the sum of income for all participants compared to the case with no punishment is lower. For  $p = 3, 4$  and higher the scenario goes toward full contribution. In this case, they registered that the sum of income for all participants, after punishment costs, increased.

### **3 The experiment**

As mentioned, I am conducting a public good game with punishment. My main purpose is to find out if individual's punishment behavior varies with respect to group size. Is free riding with respect to punishment effort more prominent in larger groups? I will explore this by two treatments with different subjects each time.

#### **3.1 *Game technicalities***

In each of the two treatments, I had 16 subjects recruited from the student mass at the University of Oslo. When showing up at the experiment location participants registered by checking off a name list, drew their placement at random, and got a paper containing rules of conduction. Afterwards they found their seat, and waited until the rest of the participants showed up.

We were three persons who conducted different experiments at the same time, so we cooperated about carrying out the experiments. We divided the areas of responsibility in such a way that for my experiment, all I did was controlling the Z-tree program. This way I had two assistants, one who read the instructions and one registered the participants at arrival. Both also helped during the experiment if needed. The same person read the instructions in each treatment.

##### **3.1.1 Set up**

When we were ready, all participants sat by their individual computer with Z-Tree started. The computers were set in two rows in such a way that the participants had their back to each other. In addition, walls were set up between each computer dividing all the participants. We did this to minimize communication and other sort of bonding that could influence their responses.

### 3.2 Rules

Every participant got a copy of the instructions at their place. Moreover, my assistant read the exact text aloud. The instructions were in Norwegian and the whole session took place in Norwegian.

During a treatment, we repeated the game ten times (periods). Each period consisted of two phases, contribution- and punishment phase. Each participant was endowed with 20 EMU (experimental monetary unit). 1 EMU is worth exactly 0.75 NOK<sup>iii</sup>, so their initial endowment was worth 15 NOK.

The first decision they had to make were how much of their endowment they were willing to contribute towards a common project, given the fact that they faced punishment in the next phase. Instructions conveyed this to the participants.

I will here write out the instructions in English (experiments were conducted in Norwegian), and show you the pictures of the screens as they showed up during a period. The screen pictures were not part of the instruction to the participants. The only difference between the two treatments was group size in the punishment phase. Written in brackets are rules applying to treatment with groupsize four, from now on defined as 16-4 (other treatment as 16-1) in punishment.

<start of instructions>

#### 3.2.1 Introduction

Welcome to this experiment and thank you for participating. Data from this session will be used in a master thesis to get knowledge of factors that influence people's decisions.

Your compensation for participating will depend on your, as well as the other person's choices.

---

<sup>iii</sup> 100 Nok = 12.35 Euro as of the 13th of April 2007

### 3.2.2 Basic rules and information

1. All answers will be collected anonymously.
  - This means that all information is anonymous. Even if you later in the experiment will get information of choices made by others, you will not be able to pinpoint who made those choices.
2. All cell phones have to be shut off.
3. No communication between participants is allowed.
4. You are not allowed to use any other program available on the computer other than the one you see in front of you now.
5. If you need help; raise your hand and we will assist you.

### 3.2.2 Setup

You are now 16 persons participating. We are doing the experiment in 10 equal periods (rounds), each consisting of two stages; one contribution stage and one punishment stage.

#### *Stage 1 (Contribution phase)*

In the beginning of each period, you all are endowed with 20 EMU (experimental monetary units). 1 EMU is worth 0.75 NOK. Of this endowment, you are to decide how much you want to contribute to a common project. The total contribution from all 16 will be multiplied by a factor of 2, and then divided equally between all 16.

Ex 1: If every of the 16 participants contributes all their endowment, the common project will get 16 times 20 EMU equal 320 EMU. This is multiplied with 2, so thus we have 640 EMU to distribute equally between all 16. I.e. each participant is left with 40 EMU after the contribution phase.

Ex 2: If all the others contribute all their endowment, and you nothing, you will get  $(15 \cdot 20 \text{ EMU} \cdot 2) / 16 = 37.5 \text{ EMU}$  from the common project. You will then be left with 57.5 EMU, and the other 15 gets 37.5 EMU each after the contribution phase.



You will then be asked to fill in your choice in the empty field and press “Contribute“.

The screenshot shows a web interface for a contribution experiment. At the top, there is a header bar with two sections. The left section is labeled "Periode" and contains the text "1 av 10". The right section is labeled "Gjenstående tid i sekunder:" and contains the text "41". Below the header, the main area is a large, empty rectangular box. In the center of this box, there is a small text label "Din beholdning i bidragsfasen" followed by the number "20". Below this, there is another text label "Ditt bidrag i bidragsfasen" followed by a small, empty rectangular input field. In the bottom right corner of the main area, there is a button labeled "Bidra". Below the main area, there is a section labeled "Informasjon" containing two lines of text: "Fyll inn ditt bidrag og trykk 'Bidra' for å fortsette til neste skjermbilde." and "Hvis flere spørsmål. Rekk opp hånden og vi vil komme rundt og hjelpe dere."

**Screen 1:** Here the participant will see his endowment (20) in the first line and in the second he can assign his contribution. This contribution must be between 0 and 20.

A new screen will appear with limited information of the outcome of stage 1.

Periode											
1 av 10	Gjenstående tid i sekunder: 29										
<table> <tbody> <tr> <td>Ditt bidrag til prosjektet</td> <td>16</td> </tr> <tr> <td>Gjennomsnittelig bidrag fra alle deltagere</td> <td>8.4</td> </tr> <tr> <td>Din beholdning etter bidrag</td> <td>4</td> </tr> <tr> <td>Din inntekt fra prosjektet</td> <td>16.9</td> </tr> <tr> <td>Din samlede inntekt etter bidragsfasen i denne perioden</td> <td>20.9</td> </tr> </tbody> </table>		Ditt bidrag til prosjektet	16	Gjennomsnittelig bidrag fra alle deltagere	8.4	Din beholdning etter bidrag	4	Din inntekt fra prosjektet	16.9	Din samlede inntekt etter bidragsfasen i denne perioden	20.9
Ditt bidrag til prosjektet	16										
Gjennomsnittelig bidrag fra alle deltagere	8.4										
Din beholdning etter bidrag	4										
Din inntekt fra prosjektet	16.9										
Din samlede inntekt etter bidragsfasen i denne perioden	20.9										
<div>Fortsett</div>											
<p>Informasjon</p> <p>Dette er en oversikt over resultater fra bidragsfasen.</p> <p>Trykk "Fortsett" når du er ferdig med å se over.</p>											

**Screen 2:** This screen shows the outcome after the contribution phase were the total contributions to the public good are multiplied by 2, and then divided equally to all 16 participants.

This information is:

1. Your contribution to the project
2. Your average contribution for all players together
3. Your income from tokens you chose not to contribute
4. Your income from the project
5. Your total income from stage 1

### Stage 2 (Punishment stage)

In the next stage, you will get an overview over all the other participants' choices (screen 3a) [your group member's choices (screen 3b)].

Periode  
1 av 10

Gjenstående tid i sekunder: 119

Den andres bidrag	Den andres inntekt (denne periode) etter bidragsfasen	Straffepoeng du vil gi til den tilhørende deltager
6	30.9	
4	32.9	
15	21.9	
5	31.9	
8	28.9	
12	24.9	
16	20.9	
13	23.9	
9	27.9	
2	34.9	
11	25.9	
1	35.9	
10	26.9	
14	22.9	
6	30.9	

Beregn

Dine resultater fra bidragsfasen  
Ditt bidrag til prosjektet 3  
Din inntekt etter bidragsfasen i denne perioden 33.9  
Gjennomsnittlig bidrag for alle deltagere 8.4

Informasjon  
Her kan du angi straffepoeng til de andre deltagerne. Du kan gi maks 10 straffepoeng til hver enkelt, og totalkostnaden for deg kan ikke overstige din beholdning etter bidragsfasen.  
1 straffepoeng koster deg 1 EV og påfører den respektive deltager en reduksjon på 3 EV i beholdning.  
Alle tomme felt må fylles ut med et tall mellom 0 og 10.  
Du må trykke "Beregn" før du kan trykke "Straff". Du har du mulighet til å se kostnad for deg og eventuelt gjøre om ditt valg.

Totalkostnad for deg av straffepoengene du gir -----

Straff

**Screen 3a:** To the left the participant sees the other 15's contributions and incomes from the contribution phase. Behind each individuals information there is an empty field for assigning punishment points.

To the right are the participants' own contribution and income. In addition, there is also the average contribution for all participants. Beneath this is a general information box about the cost and effect of punishment points assigned.

Periode		1 av 10		Gjenstående tid i sekunder: 117	
Den andres bidrag	Den andres inntekt (denne periode) etter bidragsfasen	Dine resultater fra bidragsfasen			
7	23.8	Ditt bidrag til prosjektet 2			
5	25.8	Din inntekt etter bidragsfasen i denne perioden 28.8			
5	25.8	Gjennomsnittlig bidrag for alle deltagere 5.4			
2	28.8				
10	20.8				
8	22.8				
3	27.8				
5	25.8				
12	18.8				
9	21.8				
6	24.8				
1	29.8				
		Informasjon			
		Her kan du angi straffepoeng til de andre deltagerne (i din gruppe). Du kan gi maks 10 straffepoeng til hver enkelt, og totalkostnaden for deg kan ikke overstige din beholdning etter bidragsfasen.			
		1 straffepoeng koster deg 1 EV og påfører den respektive deltager en reduksjon på 3 EV i beholdning.			
		Alle tomme felt må fylles ut med et tall mellom 0 og 10.			
		Du må trykke "Beregn" før du kan trykke "Straff". Da har du mulighet til å se kostnad for deg og eventuelt gjøre om ditt valg.			
Den andres bidrag	Den andres inntekt (denne periode) etter bidragsfasen	Totalt kostnad for deg av straffepoengene du gir -----			
4	26.8				
6	24.8				
1	29.8				
		Beregn			
		Straff			

**Screen 3b:** To the left the participant sees the other 15's contribution and income from the contribution phase. Behind each individual of your group's information, there is an empty field for assigning punishment points.

To the right are the participants own contribution and income. In addition, there is also the average contribution for all participants. Beneath this is a general information box about the cost and effect of punishment points assigned.

You will get the following data:

1. Every others contribution
2. Every others income

In addition, there will be an empty field behind each participant [in your group] where you are to assign deduction points. All this information is anonymous so it is not possible to pinpoint whose information it is. There will also be an overview of your own data plus the average contribution for all.

One deduction point assigned costs you 1 EMU. If you assign one deduction point to an other participant, his payoff will be reduced by 3 EMU. You can give up to 10 deduction points to each participant, but no more in total than your income from previous stage. All fields must

be filled out with a number between “0” and “10”. Before you can press the “Punish” button, you have to push the “Calculation” button as well. This will show you your total cost of deduction points assigned. You can change your mind, or press “Punish” for finally assigning your punishment.

After this, a new screen appears with an overview of the end-result for this period.

#### Screen 1 : End result for one period

Periode		Gjenstående tid i sekunder: 26	
1 av 10			
Din inntekt etter bidragsfasen	34.9		
Din kostnad av å dele ut straffepoeng	4		
Antall mottatte straffepoeng til deg fra andre	1		
Inntektsreduksjon for deg av straffepoeng fra andre	3		
Din inntekt etter begge fasene (bidrag- og straffefase)	27.9		
Din totale inntekt sammenlagt for alle perioder til nå	27.9		
		<b>Fortsett</b>	

**Informasjon**  
 Du kan nå se resultatene for hele denne perioden bestående av en bidragsfase og en straffefase.  
 Etter at tiden er gått ut, eller hvis alle har trykket "Fortsett", så vil eksperimentet fortsette.

This screen shows the result after one period. The next screen is then screen 1 i.e. next period starts over again with 20 EMU in endowment.

The information is:

1. Your income at the first stage
2. Your cost of assigning deduction points
3. Amount of assigned deduction points to you from others
4. Income reduction through deduction points assigned to you from others
5. Your income this period, after punishment
6. Your total income including all previous periods as well as this
7. *Your ID (only last round)*

The income each period will not be less than zero. If this happens, your income will be set to “0”. Your income from each of the 10 periods will be accumulated and this will decide your final payoff.

### 3.2.3 Questionnaire

After the ten rounds, there will be a short questionnaire for background information. This will not influence your payout or results from the experiment.

End of instructions.

<end of instructions>

The questions I asked were:

- Sex (male, female)
- Age (18-75)
- What is your field of study? (economics, mathematics, psychology, others, not student)
- On what level of education are you? (1-3 years [bachelor level], 4-5 years [master level], over 5 years, not student)
- How did you get to know about this experiment? (email, poster, flyer, friends)
- Have you participated in economic experiments before? (yes, no)
- If you are interested in participating in forthcoming experiments, please fill in your email address.
- Can you relate this experiment to situations in daily life, if so which?

## 3.3 More on game setup

The above sections were only an exact translation of instructions as presented to the participants. I will now go into details about the setup in a more mathematical fashion.

### 3.3.1 Contribution phase

The design of this common project is a standard linear public good game with  $n$  participants, each endowed with  $y$  EMU. As mentioned they then decided their contribution  $c_i$ ,

$0 < c_i < y$ , and the rest  $y - c_i$  were kept for themselves.  $i$  corresponds to participant  $i$ . The sum of all contributions from all  $n$  participants was multiplied by a factor  $\alpha$ , and then divided out equally to the  $n$  participants. Then we will get the outcome for a participant after the contribution period,  $\pi_i$ , like this.

$$(1) \pi_i = y - c_i + \frac{\alpha}{n} \sum_{i=1}^n c_i$$

The crucial part here is to create the right incentives by setting the value of  $\alpha$ . We want a unique Nash equilibrium, for preferences of homo economicus<sup>iv</sup>, where no one contributes towards the common project. For this to hold we must set  $0 < \alpha < n$ . Then it will be best for every individual not to contribute because 1 EMU contributed only returns  $\frac{\alpha}{n}$  EMU, which is less than 1 EMU. A contrast to this is that it is best for the group as a whole to contribute everything. I conveyed these points by two examples to the participants. The first were where everyone contributed all his or her endowment, the social optimum. The second all other except one, who contributed nothing, contributed all his or her endowment. I did this to enlighten the fact that you could earn more money than the others could by freeriding.

### 3.3.2 Punishment phase

The next step was the punishment phase. Here the participants got an overview of the other participant's contributions and their income from the contribution phase. With this information at hand, they could assign punishment points towards other participants, three others in 16-4 and all 15 other in 16-1.

One important point here is anonymity. During the treatments, subjects got no information on who contributed, or punished how. The only information was a list, randomized every period, containing information on all subjects' contribution and income, but no information telling who contributed what (see screen 3a and 3b). In addition we changed groups in the 16-4 treatment every period, making it impossible to recognize any pattern in individual behavior that could identify any participant.

---

<sup>iv</sup> Here, I define Homo Economicus to only have preferences of monetary outcome for himself, the more the better.

They could assign between one and ten points towards each individual, but no more in total than their income from contribution phase. In other words, for each participant they were allowed to assign punishing points,  $p_{ij}$ , where  $0 \leq p_{ij} \leq 10$  and  $\sum_{j \neq i} p_{ij} < \pi_i$ .  $p_{ij}$  is participant's  $i$  assigned punishment points towards participant  $j$  in his own group.

As well as assigning points it is obvious that each participant faced the possibility to get punishment points by others. This is denoted by  $p_{ji}$ , participant  $i$  received punishment points from participant  $j$  in his own group (15 others in 16-1 and 3 others in 1-4). One punishment point assigned cost the sender  $\beta$  EMU and inflicted  $e$  EMU reduction in the receiver's final payoff this period. In addition, no participant could get a payoff below zero in a single period. If that happened, payoff was automatically set to zero for that period. We told every participant this fact through the instructions.

The contribution- and punishment phase results in a final payoff for period  $t$  given by:

$$(2) \quad \Pi_i^t = \max \left\{ 0, \pi_i^t - \beta \sum_{i \neq j} p_{ij}^t - e \sum_{j \neq i} p_{ji}^t \right\}$$

In my treatment  $n=16$ ,  $\alpha=2$ ,  $y=20$ ,  $\beta=1$  and  $e=3$  so the final formula for my experiment will be:

$$(3) \quad \begin{aligned} \Pi_i^t &= \max \left\{ 0, y - c_i^t + \frac{\alpha}{n} \sum_{i=1}^n c_i^t - \beta \sum_{i \neq j} p_{ij}^t - e \sum_{j \neq i} p_{ji}^t \right\} \\ &= \max \left\{ 0, 20 - c_i^t + \frac{1}{8} \sum_{i=1}^{16} c_i^t - \sum_{i \neq j} p_{ij}^t - 3 \sum_{j \neq i} p_{ji}^t \right\} \end{aligned}$$

The final profit for every participant is all periods' payoffs accumulated, giving the total payoff range for one participant for the entire treatment.

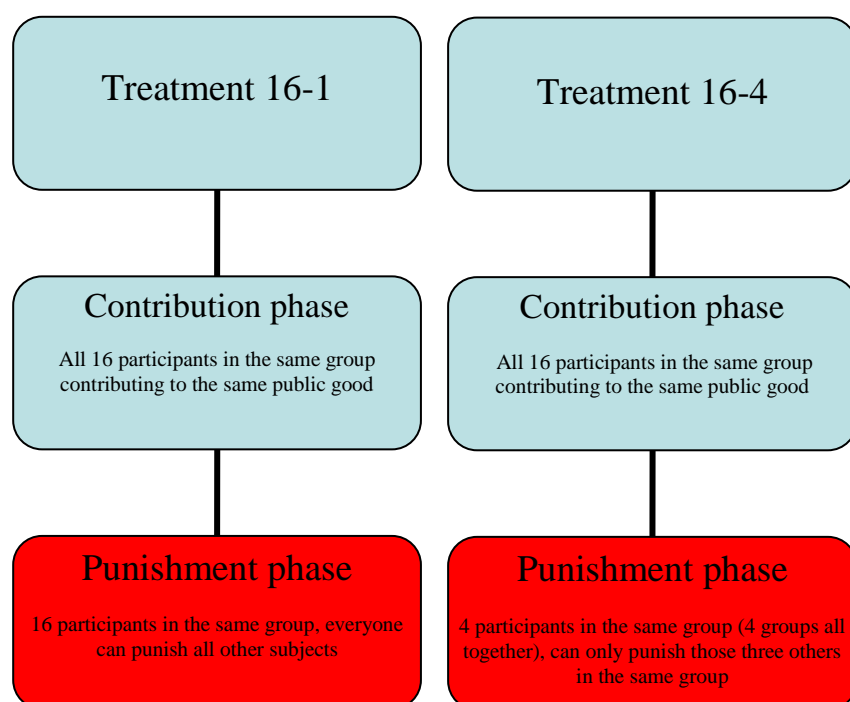
$$(4) \quad \hat{\Pi}_i = \sum_{t=1}^{10} \max \left\{ 0, y - c_i^t + \frac{\alpha}{n} \sum_{i=1}^n c_i^t - \beta \sum_{i \neq j} p_{ij}^t - e \sum_{j \neq i} p_{ji}^t \right\}_t = \sum_{t=1}^{10} \Pi_i^t$$



On top of this came the show up fee as well. 50 NOK for the first treatment, and 100 NOK for the second.

### 3.4 Treatment differences

Below is a simplified representation in the two treatments. We see that the treatments is equal in the contribution phase, and that the only difference is how many you can punish in the punishment phase.



My intention is to hold as much as possible equal in the two treatments. We see that all players contribute to the public good as a united group, implying that we probably have strong strategic incentives (reputation) for punishing in the punishment phase. Since this effect is close to identical in the two treatments, it should give low distortion of data.

After the contribution phase in my 16-4 treatment, you got an overview over all the other participants, but it was only possible to punish three randomly and exogenously selected others in your contribution group. The group size was then four, and I used the stranger function in Z-Tree i.e. the groups changed every period<sup>v</sup>. This is also an argument for the above statement about reputation building not distorting data.

<sup>v</sup> It is an option in Z-Tree called stranger. This option shuffles the group composition every period and thereby creates different groups.

In the 16-1 treatment, the contribution phase was equal to 16-4, but the punishment phase was not. In the punishment phase in 16-1 the participants got the information of all the others, as in 16-4, but now they could punish every one of the 15 other participants. This way I ensured that, I only changed one factor between the two treatments. So considering the above section, I believe I made the treatments as equal as possible in all other aspects then the group size in punishment phase.

Regarding my experiment, some papers have been looking into groupsize effects with both contribution and punishment. Jeffrey Carpenter has been doing some experiments involving different groupsize (Carpenter, 2002). In the treatment, he had two treatments of the following difference, one with groupsize of five and one with groupsize of ten. The distinction from my thesis is that the groups in Carpenter's paper were identical in both contribution and punishment phases (I only varied the size in punishment phase). His conclusion in this paper was that groupsize did not have any or little effect on punishment, but contribution rose significantly more rapidly and to a higher level in the treatment with groups of ten.

One of the most recent papers (Carpenter, 2006), discuss this topic loosely in the end by introducing shirking cliques. Shirking cliques can submerge if the team size is so large that there develops smaller working groups. If for example in such a case members of this working group can deter reciprocators (see chapter 4 for reciprocity) from punishing shirkers, we have a so-called shirking clique. However, they conclude under their assumptions that group size is not significantly influencing how or how much people punish.

## 4 Theory

In this experiment, we look into some of the fundamental human traits. It is almost like a version of crime and punishment, but with crime replaced by not contributing. How do people react to the level of contribution, prospective punishment, or punishment in it self?

In my experiment monetary incentives is a main force. I will, for that reason, in this and the next chapters discuss different preferences for monetary outcome.

### 4.1 *Homo economicus*

When we represent the economic stereotype of Homo sapiens, we often use the term Homo economicus. This is a term introduced into economics as to simplify theoretic models. Many of well-known and fundamental models, both in macro- and micro- economics, have this requirement.

One definition of Homo economicus is that he is a rational and self-interested actor<sup>vi</sup>, with preferences only for his own access to material goods. With rationality in this setting, we mean that the person maximizes utility, using efficiently the options available to him. In my experiment, an immediate characterization of a Homo economicus participant would be one who will try to get as high monetary outcome for himself without regard for any of the other participants (assuming he prefers monetary outcome).

In this experiment, we ran the treatments for 10 rounds. We can then argue that even a rational and self-interested actor may contribute and punish to build up reputation. He may want to give an impression of himself as a cooperator, by contributing, to make others contribute to the public good. He would even punish others at a cost for himself if he thinks he will gain from the increase in contribution in the following rounds. This would only hold if one or more of the players were not a utility-maximizing Homo economicus. Let me clarify. Assume all participants are rational and self-interested as by the definition of Homo economicus. Then everybody knows that in the last round no one will contribute or punish. No punishment will happen when this is costly to the punisher, as it is in this experiment. By

---

<sup>vi</sup> [Homo economicus](#) from Wikipedia, the free encyclopedia.

this, there will also be no contribution, since a Homo economicus actor will be better off by not contributing given all the other participants choices (contributions). We realize by using backward induction that we have subgame-perfect Nash equilibrium (Robert Gibbons, 1992) for all rounds that imply no contribution and no punishment. I.e. there will not be any contribution or punishment in any of the rounds.

## **4.2 Other theories**

From prior experiments mentioned in the background chapter, we know that people do cooperate and punish others. We need other models to portray human behavior and I will explain some of them and tell why this is relevant for my experiment.

### **4.2.1 Inequity aversion**

People tend to compare themselves with others quite extensively. An inequity adverse person has preference for equality. He prefers equality when comparing himself with others, and any inequality reduces his utility. This does not mean that he does not care for monetary payoff for himself, indeed he does. For example if you and another person do the same type of work, contributing the exact same effort, have the same pay, but suddenly one day you get a raise. If you are an inequity averse person, you will not enjoy the full increase in utility due to this extra money. The inequity between you and your coworker is reducing your utility. Reduction will not be so severe as to offset the whole gain in utility (with normal parameters), but a reduction nevertheless.

This of course raises new difficult questions on how people measure and perceive fairness of outcomes (Fehr, Schmidt, 1999). In this paper, Fehr & Schmidt presented an intuitive model to represent utility of individuals in such a setting with equity and fairness. A requirement for this model is that people have information about others' payoffs, as they will have in my experiment. The base model is that the subject compares his own outcome with the others'. If his own outcome is less, or higher, than that of the other person, his utility will be lower. An equation for this statement is as follows.

$$(5) \quad U_i(x) = x_i - \alpha_i \max |x_j - x_i, 0| - \beta_i \max |x_i - x_j, 0| \quad \text{where } i \neq j$$

Here is  $U_i(x)$  individual  $i$ 's utility that consists of three parts.  $x_i$  is his monetary payoff and then if individual  $j$  has a higher payoff, he subtracts a share,  $\alpha_i$ , of this difference,  $x_j - x_i$ . If however individual  $i$  has a higher payoff, then he will subtract a share  $\beta_i$  of this difference,  $x_j - x_i$ . Typically we have that  $\alpha_i > \beta_i$ , i.e. he weighs inequity in his disfavor more than if it's in his favor.

Equation (5) is for two individuals and is derived from a more general function (see the Fehr & Schmidt article), but I believe the two participant equation expresses the experimental situation under punishment phase adequately. If you remember that participants got an overview of information of all the others' individual choices, it should be natural to compare oneself with one participant at a time. This way, if the other participant had contributed little and gained much from the public good compared to oneself, the incentive to punish should be clear. Another possibility is that he (in addition) looks at the average contribution for all and compares this with both his own and the other person's contribution.

Example: If you contribute 18 EMU and the other person only 2 EMU in a two player public good game. Then you will have 22 EMU and he will have 38 EMU after the contribution phase. Inserted in (5) we get:

$$\begin{aligned} U_i(x) &= 22 - \alpha_i \max |38 - 22, 0| - \beta_i \max |22 - 38, 0| \\ &= 22 - \alpha_i 16 \end{aligned}$$

Then he has incentive to punish as to minimize the difference in payoff.

However, the problem for the participant is not only if he is going to punish, but also how much to punish. In this experiment, he has to take into account all the other participants actions. How much will they punish? Will they punish me? That is why groupsize can be of importance. With many potential punishers who can punish each other it may be the case that those who punish believe that someone else will step up and punish defectors. In a way, they want to freeride on others punishments. A smaller group there a fewer others who can punish, so you will feel a stronger incentive to step up and punish.

In the above statement, I linked inequity-aversion with punishment. When one period is over the participants have gained experience of how the others contributed, as well as punished.

For example if you realize that you gave less than a majority (or average) of the others (this requires an inequity model with  $n$  subjects, I will not go into that one here). If you expect others to behave approximately the same way in the next round, and if you consider inequity of each round separately, you will contribute by inequity-aversion more in the next round, given no punishment. Alternatively, maybe the other way around is more evident. When you gave more than the average for all in one period, you will try to even out this inequity and contribute less the next period, still assuming no punishment. This way you hope to get closer to the average contribution and thereby trying to maximize your utility.

### 4.2.2 Reciprocity

Reciprocity is a preference to repay kindness with kindness and unkindness with unkindness. As opposed to inequity-aversion, reciprocity takes into account the process for the outcome (Sobel, J., 2005). Positively interpreted actions elicit positive responses and vice versa. This opens up for a more subjective and complex thought, and the model reflects that. I will not go into mathematical details, as it is more complicated than the one above. When we are dealing with models of reciprocity, we need to assess the beliefs of the involved subjects.

**Example:** Let us get back to our study group mentioned in chapter 1. You suspect that a person in the group did not prepare adequately, and that he did this as an unkind act. This reduces your utility (you feel bad or maybe some anger). You decide to sanction against him by telling an insulting joke, frowning, or telling him straight up. By doing this you feel a bit better. Your utility has increased. Not fully, but you feel better than if you did not take action.

The effect of reciprocity on contribution and punishment are similar to those in inequity-aversion. After the contribution phase you may prefer to punish those you believe have contributed little, believing they are unkind. You then feel better by punishing as stated above. Another reaction takes place for the next round's contribution. Let us say you feel that others have contributed less than you because they are unkind. Then you contribute considerably less the next round as to impose unkindness towards the other participants. We see that the effect is similar, but the psychological cause is different. In inequity-aversion, we punish to adjust

outcome as to make people equal, while in reciprocity we react more directly by feelings of fairness.

### 4.2.2 Altruism

In economics several versions of altruism have been, and still are, discussed. We start with the classical approach where a person's utility depends on others utility. I.e. I get higher utility when others' utility increases.

There are two cases of special interest, pure- and impure altruism. Pure altruism, in this thesis's context, is that a person cares about others' access to a public good. In this experimental setting this can manifest itself by higher contribution toward a public good. Worth mentioning is the fact that pure altruists also contribute more even when others contribute less.

$$(6) \quad U_i = V(x_i, G)$$

Here  $x_i$  is own payoff and  $G$  is the magnitude of the public good. Both first derivatives are positive. We see then with proper assumptions we can get some contribution based on pure altruism (Andreoni, 1998). Andreoni has also published some papers containing discussion of impure altruism. This type of altruism builds on the pure model, but in addition, there is a factor of "warm glow" by giving. A person gets higher utility by just giving.

$$(7) \quad U_i = W(x_i, G, g_i)$$

Here  $g_i$  is what he gives, for example to charity (Andreoni, 1990). In a public good game, we can think of this as if the participant gets higher utility by contributing both because he raises the value of the common good, and because he values his own contribution in it self.

When it comes to punishment, Fehr and Gächter looked into why people punished each other even though there was no monetary gain in doing so. They did this with a public good game with punishment. They repeated the game, but rearranged the groups for each period such as no participants ever were on the same group twice. There were no strategic incentives to

explain their punishment behavior, so they concluded that punishment was altruistic (Fehr & Gächter, 2002). It can be strange to think about punishment as an altruistic act since both the punisher and punished get reduction in income, but remember that this experiment was repeated. They interpreted the punishment altruistic in the sense as they were willing to take the cost of punishing as to uphold a large contribution, even though they knew they never would meet that person later in the experiment.

In my experiment, every participant contributed to the same public good each round. One important factor of punishment is therefore an incentive to build reputation, but the effect of altruistic punishment is probably present.



## 5 Results

The treatment with 16 participants and punishment groupsize of four went well. Unfortunately the treatment where all has the ability to punish all treatments crashed. I put much effort in the programming and preparation for the experiments, so I did not have time to redo the experiment which failed. What I got of data is ten full rounds of the successful (16-4) treatment and three of the failed one (16-1). I did two additional rounds, and will be using these as to get five rounds of 16-1.

### 5.1 Problems

The treatment where there was only the group who could punish each other went without trouble at all (16-4). The other treatment however did not go as smoothly. It stopped after few rounds the first time we ran it with participants, and almost immediately after start, some computers dropped out of the Z-tree client's overview<sup>vii</sup>. At that time, we did not have the skills to tackle the problem satisfactory, so we cancelled the experiment and paid every participant their show up fee.

Immediately after the participants had left, we tested the experiment for us selves in full scale, i.e. all 16 computers connected. This time we made ten satisfactory rounds, and everything seemed to be in order as it had been in all the other full-scale test experiments.

For that reason, I did not change any major lines in my program and conducted a second experiment with the same program (minor changes). However, we changed the placement of the router, as explained later. After three successful rounds, it stopped again. This time I got the data from the three first rounds, so there is some data to analyze. We decided to restart the experimenter-computer and just run it for two more rounds. These two rounds went by without trouble. The participants then got the payoff from these two rounds multiplied by 5, as we informed them. We did this to get to the questionnaire part of the experiment (can only be started after successful ending of a treatment).

Under follows some solutions and notes about the problem.

---

<sup>vii</sup> This is a table over computers connected to the Z-Tree experiment. This is available for the experiment leader in realtime during a session.

### 5.1.1 Router

Sometimes when computers fell out of the clients table the reason were that the client computer (the computer the participant used) no longer were connected to the router. In the first experiment, the router stood in one end of an oblong room. With 16 participants, there may be a possibility that so much mass (bodies, walls, etc.) would block or at least hinder the communication between router and computer. A support for this theory is that computers at the far end of the router most commonly lost connection. The next time we placed the router at the middle of one of the long walls, and as far as we are aware of we did not experience any more network dropouts. Therefore, it can be an idea to have a more powerful router and think well about the placement of this.

### 5.1.2 Program

As mentioned over I do not think there is a direct fault in my programming. A reason for this is that I have run the program in full scale many times at my own computer as well as when testing the lab. Other limitation that can be the cause is the size of the contracts table.

Contracts table is a special table in Z-Tree that collects data from every sender (the one who assigns punishment points) to every receiver (participants who receives the punishment points assigned from sender). With 16 participants punishing 15 others, this table has 240 lines with data. The other treatment with groups of four has only 48 lines of data. That this table is updated live as soon as someone enters a number at their screen is another point. This way I believe that there were too much going on at the same time for the network to handle. In our trials, we were only four to six persons doing the testing. It was impossible for us to enter data at such a rate to get this kind of error, and that one of the reason we did not discover this fault sooner.

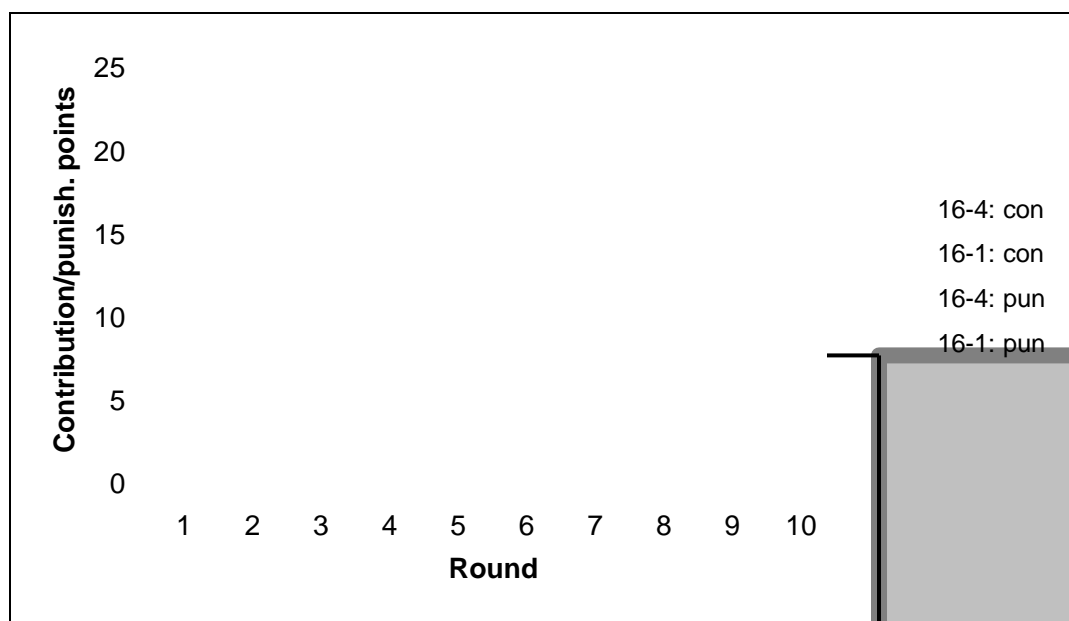
Another point regarding this is that the program locked when I were adjusting the width of tables at the server computer. It may be the case that this is what caused the treatment to shut down. One solution to this is to not do anything on server/experimenter computer during treatment, this way ensuring that all computer resources is used for handling dataflow.

## 5.1 Data summary

The first table I will present is average numbers for contribution and punishment for all rounds. I will characterize round four and five in 16-1 as not entirely consistent, when the treatment crashed after round three, but still use them in my analysis.

### 5.1.1 Contribution and punishment

First, we look at the development of contribution and punishment in the two treatments. In papers mentioned in the background chapter there were substantial evidence that with my cost/effect ratio contribution should rise through the rounds. My experiment confirms this, by the graph below. Contribution increases steadily, but more rapidly in the 16-1 treatment. In addition, it looks from my data that in the 16-1 treatment, contribution is at a higher level compared to 16-4.



**Graph 1:** Average contribution and punishment points assigned each round for both treatments.

When we look at punishment, we see that there is quite extensive punishment in both treatments for the first round. However, it seems as it stabilizes for both treatments. Another point is that it looks as if 16-1 punishment is lower than 16-4 punishment, indicating support for my hypothesis. As discussed below, contribution has a great impact on punishment, so if we look at punishment relatively to contribution we have an effect that defies my hypothesis. Since contribution seems to be higher in 16-1 than in 16-4, this will have an impact on how

people punish, and in this setting weakening my hypothesis. These effects are controlled for in my regression.

### 5.1.2 Individual punishment received

We have seen above that average contribution and punishment points assigned differs in the two treatments. The problem however is to test my main hypothesis.

**Hypothesis 1:** For any given individual contribution by individual  $i$ , individual  $i$  is punished less by others in the 16-1 treatment than in the 16-4 treatment.

I will do that by linear regression, testing for

$H_0$ : For any given individual contribution by individual  $i$ , individual  $i$  is punished equal by others in the 16-1 treatment and in the 16-4 treatment.

$H_1$ : For any given individual contribution by individual  $i$ , individual  $i$  is punished differently by others in the 16-1 treatment than in the 16-4 treatment.

To test this properly I have to find the variables that influence punishment points received the most. I will here list up those I believe are essential:

- “group”: this indicates which experiment the data is from. “0” is 16-4 and “1” is 16-1. “group” will be the variable we will check for influence on received punishment points.
- “contribution”: this is the contribution toward the public good of the participant receiving punishment points. Obviously, this should have a impact on punishment points received. High contribution will probably give less punishment points from others.
- “receivedpoints”: total numbers of punishment points received by the participant from the other potential punishers. This is our dependent variable, making the two others independent.

From table 1 we see that received points correlates quite little with group(treatment). The same conclusion for contribution versus group, but we see that the correlation between contribution and received points is quite high, as anticipated.

	group	contri~n	receiv~s
group	1.0000		
contribution	0.0722	1.0000	
receivedpo~s	-0.0817	-0.7768	1.0000

**Table 1:** Correlation matrix for selected variables.

We are now ready to do a regression with received points as the dependent variable. The regression formula is as follows:

$$(8) \text{ recievedpoints} = \text{constant} + \alpha \cdot \text{group} + \beta \cdot \text{contribution}$$

The result is represented in table 2:

Source	SS	df	MS			
Model	4017.36626	2	2008.68313	Number of obs =	240	
Residual	2632.92957	237	11.1094075	F( 2, 237) =	180.81	
Total	6650.29583	239	27.8255056	Prob > F =	0.0000	
				R-squared =	0.6041	
				Adj R-squared =	0.6007	
				Root MSE =	3.3331	

receivedpo~s	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
group	-.2879078	.457594	-0.63	0.530	-1.189379	.6135633
contribution	-.648927	.034315	-18.91	0.000	-.7165284	-.5813255
_cons	13.33072	.8144468	16.37	0.000	11.72624	14.9352

**Table 2:** regression results

We see that “contribution” has an impact on “recievedpoints” which is negative and significantly different from zero. The negative impact means that a high contribution leads to less punishment points received from others. “group” has also a negative impact on “recievedpoints” (supports my prediction), but we see that the p-value is way too high for the effect to be significantly different from zero. We cannot discard the  $H_0$ -hypothesis based on the data produced in my experiment. However, I believe that this opens up for additional experiments to get a complete dataset to analyze.

## **5.2 Why this “simple” regression?**

It is important to notice that I look at total punishment points received and not punishment points assigned towards each individual. This is because I needed consistency in my data. If I had looked at assigned punishment points towards each individual, I would clearly created additional difficulties when you in one treatment punish three participants and in the other fifteen. Obviously, you punish each individual more when there is only three to punish compared to the treatment with fifteen and it would be unnecessarily complex to design a regression taking this into account.

### **5.2.1 Critique**

One main weak point with this regression is that the participants learn from round to round. In many papers, they take the four first rounds as learning rounds. I.e. four first rounds are not taken into account in the analysis. This is to ensure that the subjects have a feel for the mechanism or rules of the experiment. Since I did not accomplish to complete one of the treatments, I did not have the opportunity to leave out any rounds.

Another aspect about learning is the fact that the outcome from the previous rounds has an impact on choices made in this round. When you have gotten harsh punishment because you contributed little in one round, this will probably influence your contribution in the next and following rounds. This can influence how people behave and characterize individual participants. Which types of other participants they are playing against can also influence this. Example, a low contributor can punish those who contribute little, a high contributor may not punish anybody etc. This may lead to differences in how the game is played out. I have, however not gone into this.

## 6 Summary and conclusion

Behavioral economics is about how people really behave contrary to profit maximizing actors. I have conducted a public good experiment with the opportunity to punish. The game consisted of two phases, one contribution phase and one punishment phase. In contribution phase participants chose how much to invest towards a public good. In the punishment phase, the participants could react to the others contribution by punishing individuals. The result does not lead to any conclusion since the p-value is too high. Even if the sign is negative as predicted, there is no evidence in the dataset confirming my hypothesis.

After all this work with programming and planning of the experiments, it is somewhat sad that the outcome were so limited in hard data. Nevertheless, I have gained experience in exactly those aspects of running an experiment, planning and preparations. When it comes to contribution to the behavioral economics community, I believe this experiment have some valuable insights.

First, the script is ready for further experimenting on this topic. Even though the treatment stopped because of the earlier mentioned problems, I believe it is possible to redo the experiment successfully with existing script and program solution (z-Tree). By considering the discussion in chapter 5, it should be possible to get data for the 16-1 treatment. One of the most important and easiest changes is to let the experiment run without doing any unnecessary work on the server computer.

Second, looking at the data and regression output we see that there is a negative, but insignificant effect for groupsize. However, since the size of the estimate is relatively large and with the sign predicted by my hypothesis, this seems worth looking into with further experiments.

## 7 Appendix

### 7.1 Instructions for the participants in Norwegian

Norwegian version of instructions to the participants, read aloud by one of the assistants.

#### 7.1.2 Treatment 16-1

<start of instructions>

##### Introduksjon

Velkommen til dette eksperimentet og takk for din deltagelse. Data fra dette partiet vil bli brukt i en masteroppgave for å forstå mer om hvilke faktorer som påvirker menneskers valg. Kompensasjonen for din deltagelse her avhenger ikke bare av dine valg, men også de andre deltageres valg.

##### Regler og informasjon

1. Alle valg vil bli samlet anonymt
2. Det betyr at man verken under eksperimentets gang eller etterpå vil ha mulighet til å finne ut hvem som svarte hva.
3. Alle mobiltelefoner må være skrudd av.
4. Ingen kommunikasjon mellom deltagere er lov.
5. Det er ikke lov å bruke annen programvare enn den dere ser foran dere nå.
6. Hvis du trenger hjelp; rekk opp hånden, og vi vil komme og hjelpe deg.

##### Oppsett

Dere er nå 16 deltagere. Vi skal gjøre eksperimenter i 10 like perioder. Hver periode består av to faser: En bidragsfase og en straffefase.

##### *Fase 1 (Bidragsfase)*

I begynnelsen av hver periode vil du bli tildelt en beholdning på 20 EV (eksperiment valuta). 1 EV er verdt 0.75 kr. Av denne beholdningen skal du bestemme deg for hvor mye du vil gi til et felles prosjekt, og hvor mye du vil beholde selv. Om du vil bidra, og hvor mye, er helt opp til deg. De samlede bidragene til fellesprosjektet fra alle 16 blir ganget med 2, og så delt likt ut igjen til alle 16.



Eks 1: Om alle 16 gir hele sin beholdning vil det si at fellesprosjektet får 16 ganget med 20 EV lik 320 EV. Dette ganges med 2. Altså har vi 640 EV å dele likt ut på alle 16, som gjør at hver og en av dere sitter igjen med 40 EV etter bidragsfasen.

Eks 2: Om alle gir hele sin beholdning, og du ikke gir noe. Så vil du få  $(15 \cdot 20 \text{ EV} \cdot 2) / 16 = 37,5 \text{ EV}$  igjen fra fellesprosjektet. Du vil da sitte igjen med 57,5 EV, mens de andre vil ha 37,5 EV etter bidragsfasen.

Etter at du har bestemt hvor mye du vil bidra med, trykker du "Bidra" -knappen. En ny skjerm vil komme frem med informasjon om utfallet av bidragsfasen. Informasjonen er:

1. Ditt bidrag til prosjektet
2. Gjennomsnittlig bidrag fra alle deltagere
3. Din beholdning etter bidrag
4. Din inntekt fra prosjektet
5. Din samlede inntekt etter bidragsfasen i denne perioden

Etter at du har sett over denne siden, trykk "Fortsett" -knappen for å komme til straffefasen.

### *Fase 2 (Straffefase)*

I denne fasen vil du få en oversikt over alle de andre deltagernes utfall og valg. Du vil få følgende informasjon:

1. Den andres beholdning etter bidragsfasen
2. Den andres bidrag

All denne informasjonen er anonym, og det er ikke mulig å se hvem som har valgt hva.

Det vil også være en oversikt over ditt eget utfall fra bidragsfasen

I tillegg vil det være et tomt felt bak hver av deltagerne hvor du kan angi straffepoeng.

Hvert straffepoeng du gir til en deltager koster deg 1 EV. Hvis du gir annen deltager ett straffepoeng vil dette føre til en gevinstreduksjon på 3 EV hos denne deltageren. Du kan gi opp til 10 straffepoeng til hver deltager, men ikke mer totalt enn din beholdning etter bidragsfasen.

Alle felt må fylles ut med et nummer mellom "0" og "10". Før du kan trykke "Straff" -knappen, må du trykke "Beregn" -knappen. Denne vil vise deg total kostnad for deg av de straffepoengene du har gitt til andre. Du har da muligheten til å endre ditt valg, eller fortsette til neste skjermbilde ved å trykke "Straff" -knappen.

Du vil da få følgende oversikt over resultatet for denne runden:

1. Din inntekt etter bidragsfasen

2. Din kostnad av å dele ut straffepoeng
3. Antall mottatte straffepoeng til deg fra andre
4. Inntektsreduksjon for deg av straffepoeng fra andre
5. Din inntekt etter begge fasene (bidrag - og straffefase)
6. Din totale inntekt sammenlagt for alle perioder til nå
7. *Ditt ID-nummer*

Beholdning etter endt periode vil alltid bli høyere enn eller lik null. Om resultatet skulle bli negativt vil dette automatisk gjøres om til null. Inntekten fra alle rundene vil bli lagt sammen, og dette vil avgjøre din endelige utbetaling.

Viktig: I runde 10 (den siste) vil du få en ekstra linje med informasjon (7 i kursiv over). Noter dette nummeret på utbetalingsskjema og skriv inn "1" i det tomme feltet. Dette er bare en hindring for at du ikke skal klikke deg videre før du har skrevet ID-nummeret ned.

### **Spørreskjema**

Etter den siste runden vil det bli et spørreskjema for bakgrunnsinformasjon. Dette vil ikke ha innvirkning på verken eksperimentet eller din utbetaling.

<end of instructions>

### 7.1.3 Treatment 16-4

<start of instructions>

#### Introduksjon

Velkommen til dette eksperimentet og takk for din deltagelse. Data fra dette partiet vil bli brukt i en masteroppgave for å forstå mer om hvilke faktorer som påvirker menneskers valg. Kompensasjonen for din deltagelse her avhenger ikke bare av dine valg, men også de andre deltageres valg.

#### Regler og informasjon

1. Alle valg vil bli samlet anonymt
  - a. Det betyr at man verken under eksperimentets gang eller etterpå vil ha mulighet til å finne ut hvem som svarte hva.
2. Alle mobiltelefoner må være skrudd av.
3. Ingen kommunikasjon mellom deltagere er lov.
4. Det er ikke lov å bruke annen programvare enn den dere ser foran dere nå.
5. Hvis du trenger hjelp; rekk opp hånden, og vi vil komme og hjelpe deg.

#### Oppsett

Dere er nå 16 deltagere. Vi skal gjøre eksperimenter i 10 like perioder. Hver periode består av to faser: En bidragsfase og en straffefase.

#### *Fase 1 (Bidragsfase)*

I begynnelsen av hver periode vil du bli tildelt en beholdning på 20 EV (eksperiment valuta). 1 EV er verdt 0.75 kr. Av denne beholdningen skal du bestemme deg for hvor mye du vil gi til et felles prosjekt, og hvor mye du vil beholde selv. Om du vil bidra, og hvor mye, er helt opp til deg. De samlede bidragene til fellesprosjektet fra alle 16 blir ganget med 2, og så delt likt ut igjen til alle 16.

Eks 1: Om alle 16 gir hele sin beholdning vil det si at fellesprosjektet får 16 ganget med 20 EV lik 320 EV. Dette ganges med 2. Altså har vi 640 EV å dele likt ut på alle 16, som gjør at hver og en av dere sitter igjen med 40 EV etter bidragsfasen.

Eks 2: Om alle gir hele sin beholdning, og du ikke gir noe. Så vil du få  $(15 \cdot 20 \text{ EV} \cdot 2) / 16 = 37,5 \text{ EV}$  igjen fra fellesprosjektet. Du vil da sitte igjen med 57,5 EV, mens de andre vil ha 37,5 EV etter bidragsfasen.

Etter at du har bestemt hvor mye du vil bidra med, trykker du "Bidra" -knappen. En ny skjerm vil komme frem med informasjon om utfallet av bidragsfasen. Informasjonen er:

1. Ditt bidrag til prosjektet
2. Gjennomsnittlig bidrag fra alle deltagere
3. Din beholdning etter bidrag
4. Din inntekt fra prosjektet

## 5. Din samlede inntekt etter bidragsfasen i denne perioden

Etter at du har sett over denne siden, trykk "Fortsett" –knappen for å komme til straffefasen.

### *Fase 2 (Straffefase)*

Dere blir delt inn i grupper på 4 og 4. Gruppenes sammensetning endres hver periode. I denne fasen vil du få en oversikt over alle deltagernes utfall og valg. Du vil få følgende informasjon:

1. Den andres beholdning etter bidragsfasen
2. Den andres bidrag

All denne informasjonen er anonym, og det er ikke mulig å se hvem som har valgt hva.

Det vil også være en oversikt over ditt eget utfall fra bidragsfasen

I tillegg vil det være et tomt felt bak hver av deltagerne i din gruppe (3 tomme felt) hvor du kan angi straffepoeng. Du kan altså bare straffe de som er i din gruppe, og de kan bare straffe deg.

Hvert straffepoeng du gir til en deltager koster deg 1 EV. Hvis du gir annen deltager ett straffepoeng vil dette føre til en gevinstreduksjon på 3 EV hos denne deltageren. Du kan gi opp til 10 straffepoeng til hver deltager, men ikke mer totalt enn din beholdning etter bidragsfasen.

Alle felt må fylles ut med et nummer mellom "0" og "10". Før du kan trykke "Straff" –knappen, må du trykke "Beregn" –knappen. Denne vil vise deg total kostnad for deg av de straffepoengene du har gitt til andre. Du har da muligheten til å endre ditt valg, eller fortsette til neste skjermbilde ved å trykke "Straff" –knappen.

Du vil da få følgende oversikt over resultatet for denne runden:

1. Din inntekt etter bidragsfasen
2. Din kostnad av å dele ut straffepoeng
3. Antall mottatte straffepoeng til deg fra andre
4. Inntektsreduksjon for deg av straffepoeng fra andre
5. Din inntekt etter begge fasene (bidrag - og straffefase)
6. Din totale inntekt sammenlagt for alle perioder til nå
7. *Ditt ID-nummer*

Beholdning etter endt periode vil alltid bli høyere enn eller lik null. Om resultatet skulle bli negativt vil dette automatisk gjøres om til null. Inntekten fra alle rundene vil bli lagt sammen, og dette vil avgjøre din endelige utbetaling.

**Viktig:** I runde 10 (den siste) vil du få en ekstra linje med informasjon (7 i kursiv over). Noter dette nummeret på utbetalingsskjema og skriv inn "1" i det tomme feltet. Dette er bare en hindring for at du ikke skal klikke deg videre før du har skrevet ID-nummeret ned.

**Spørreskjema**

Etter den siste runden vil det bli et spørreskjema for bakgrunnsinformasjon. Dette vil ikke ha innvirkning på verken eksperimentet eller din utbetaling.

<end of instructions>

## 8 References

- Colin Camerer, Richard H. Thaler. "Anomalies: Ultimatums, dictators and manners". *The Journal of Economic Perspectives*, Vol. 9, No. 2. (Spring, 1995), pp. 209-219
- Elinor Ostrom, James Walker, Roy Gardner. "Covenants with and without sword: Self governance is possible". *The American political science review*. Vol. 86, No. 2 (Jun., 1992), pp. 404-417
- Ernst Fehr, Klaus M. Schmidt. "A theory of fairness, competition and cooperation" *The Quarterly Journal of Economics*, Vol. 114, No. 3. (Aug., 1999), pp. 817-868.
- Ernst Fehr, Simon Gächter, (2002). "Altruistic punishment in humans" *Nature*. Vol. 415
- Ernst Fehr, Simon Gächter. "Cooperation and Punishment in Public Goods Experiments" *American Economic Review*, 2000, Vol. 90 (4); 980-994.
- Garrett Hardin. "The Tragedy of the Commons". *Science*, 162 (Dec., 1968), pp. 1243-1248
- Gerald Marwell, Ruth E. Ames. "Economists free ride, does anyone else?" *Journal of public economics*. Vol. 15 (1981), pp. 295-310
- James Andreoni, (1990). "Impure altruism and donations to public goods: A theory of warm-glow giving" *The Economic Journal*, Vol. 100, No. 401. (Jun., 1990), pp. 464-477.
- James Andreoni, (1998). "Toward a Theory of Charitable Fund-Raising" *Journal of Political Economy*, vol. 106, no. 6.
- Jeffrey Carpenter, Samuel Bowles, Herbert Gintis. "Mutual monitoring in teams: theory and experientialevidence on the importance of reciprocity" *IZA discussion paper No. 2106*
- Jeffrey Carpenter, 2002. "Punishing Free-Riders: how group size affects mutual monitoring and
- John O. Ledyard (1995) "Public Goods: A survey of experimental research" *The Handbook of Experimental Economics* edited by Roth and Kagel pp. 111-194
- Joseph E. Stiglitz, Carl E. Walsh, (2002) "Principles of microeconomics, 3rd ed." *Learning book* pp. 229
- Madan M. Pillutla, J. Keith Murnighan. "Unfairness, Anger, and Spite: Emotional Rejections of Ultimatum Offers" *Organizational behavior and human decision processes*. Vol. 68, No. 3. (Dec., 1996), pp. 208-224
- Matthew Rabin, 2002. "A perspective on psychology and economics" *European Economic Review* 46 (2002) 657 – 685
- Nikos Nikiforakis, Hans-Theo Normann. "A Comparative Statics Analysis of Punishment in
- Public-Good Experiments" *Royal Holloway, University of London: Discussion Papers in Economics* with number 05/07.
- Robert Gibbons. "A primer in game theory" *Learning book* (1992) pp. 95
- Sobel, J. "Interdependent Preferences and Reciprocity", *Journal of Economic Literature* 43, 392-436.
- the provision of public goods" *Middlebury college economic discussion paper No. 02-06*
- Urs Fischbacher (2007): *z-Tree: Zurich Toolbox for Ready-made Economic Experiments*, *Experimental Economics* 10(2), 171-178.